A METHOD OF ASSESSMENT OF A PHYSICIAN’S COMPETENCY  
BY HIS PERFORMANCE  
Proposal of a Research Project  

Background  

The past few months have seen rapidly growing acceptance of the concept of recertification as a desirable approach to demonstration of physicians’ accountability. Most of the twenty-two specialty boards have expressed formal approval of voluntary, periodic re-examination for the purpose of assuring clinical competence. One board requires periodic recertification. Four boards have set dates for initiation of recertification programs.

In addition, intensive study is being given the entire system of medical licensure and to possible approaches to periodic relicensure. Among the proposals being seriously considered is that of basing relicensure upon specialty board recertification. It seems apparent that whatever approach or approaches are ultimately adopted, they will be greatly influenced by the techniques used in the recertification process.

These rapidly developing events have focused upon the overriding question, "WHAT IS THE REAL PURPOSE OF RECERTIFICATION?"

There would seem to be general agreement that the answer is "To assure the consumer and the profession with validity the clinical competency of the physician." However, the answer to the question, "What is the most valid technique that can be devised to accomplish this goal?" has not yet been determined. The American Society of Internal Medicine (ASIM) is interested in pursuing this question to determine if it can be answered.

ASIM 4.3 a federation of 51 state component societies having a membership of over 12,000 internists in private practice and academic medicine. The American Society of Internal Medicine is an Internal Revenue Service Code 501 (c) (6) organization. ASIM has an affiliated Socio-Economic Research and Education Foundation which has been designated as a "Public 501 (c) (3)" Foundation by the Internal Revenue Service.

The 1965 ASIM House of Delegates expressed the following to be among the primary objectives of the Society:

1. To develop and maintain standards for the practice of internal medicine consistent with high quality medical care; and

2. To assure the continuing professional competence of the internist in the practice of his specialty.

In pursuit of these goals, many of ASIM’s activities over the past ten years have been directed toward developing, testing and encouraging
the acceptance of effective peer review methods as a valid approach to public accountability. The Society has published manuals in peer review and in Professional Standards Review Organization development through peer review methods. ASIM and the American Medical Association cosponsored the first major national conference for dissemination of information on peer review. Valuable basic experience in the under-explored field of office practice quality evaluation was gained in the ASIM project funded ($148,555.00) through a contract with the Department of Health, Education and Welfare.

ASIM participates in the Kellogg Foundation-funded Private Initiative in PSRO with the American College of Physicians, American Hospital Association, American Medical Association and the American Association of Foundations for Medical Care. ASIM has representation on the Board of Regents of the Institute on Professional Standards which administers an HEW contract for training personnel for PSRO's.

The Problem

Two basic approaches to the assessment of a physician's competence are now under consideration:

1. Evaluation of his knowledge by examination.

2. Evaluation of the quality of the care provided in his daily practice of medicine.

1. Evaluation of knowledge. The formal examination based on recall of knowledge comprises all or part of the recertification process plans of all Boards at the present time. The first of these, that of the American Board of Internal Medicine, offered in October 1974, is designed to make maximum use of the popular American College of Physicians Medical Knowledge Self-Assessment Program. Particular attention is given to the practical clinical applications of the material addressed by the recertification exam itself, in order to counter the frequently expressed criticism that self-assessment tests have emphasized knowledge in areas of little use to the practicing physician.

Only the American Board of Family Practice presently requires mandatory recertification, to be accomplished through a written examination every six years, 300 hours of continuing education, and a review of patient records selected by the member.

Included in the plans of several of the Boards, in addition to the traditional written examinations, is the use of computer assisted patient management simulation tests whose purpose is to display the decision-making logic processes of the testee.

Proponents of the written examination, while readily admitting the inherent weaknesses of present testing techniques, point out its similarity in form to initial certification. They
believe that it will provide an impetus to the physician's continuing education activities, since knowledge is an important and necessary component in competence. While acknowledging the importance of evaluation of practice competence, they point out that, as yet, no wholly satisfactory technique has been developed.

Opposition to recertification solely by examination is based on the following concerns:

A. The real issue of recertification (and perhaps eventual relicensure) is the actual quality of care provided.

B. There is little evidence that any examination based on knowledge recall can demonstrate competency. The self-defined limits of a physician's day-to-day practice further complicate the examination approach.

C. Knowledge is only one component of competency; other important elements include cognitive skills, interpersonal relationships, efficiency, adequacy of laboratory facilities and availability are highly important components.

D. There is considerable evidence that many (perhaps most) of the deficiencies in care detected by peer review methods are not the result of insufficient knowledge.

E. There is considerable doubt that patient care simulation tests accurately display the management decisions of the office setting.

F. Undisputed adoption of the examination approach may establish precedent over and preempt use of more valid methods as such methods are developed in the future.

To summarize it is only necessary to quote C. Barber Mueller's letter to the Editor of the New England Journal of Medicine on June 17, 1971.

*Whenever reexamination, relicensure or recertification is mentioned among medical practitioners and educators, the ensuing conversation invariably proceeds in a manner that assumes that the reexamination process will consist of a combination of continuing education courses followed by periodic examinations conducted in a fashion not unlike that of Part III of the National Boards or some of the specialty-board examinations. These examinations would thus involve simulated patient situations, simulated problem solving, a computer-based examination, a series of multiple-choice questions and oral or written examinations. I contend that such examination practices are inappropriate tools for continued assessment of a physician in practice and are not able to provide the clarity or rationale required to assess continuing physician performance in the line of duty. It is a physician's performance 'at work in a responsible setting' that must be assessed. Performance at work (what a man does), rather than performance in
2. Assessment by Performance. Recognizing the need for more objective documentation of physician competence and the likelihood of eventual recertification and possibly relicensure, ASIM's Committee on Quality Evaluation and its Task Force on Assessment by Performance, in early 1974 made a detailed study of possible evaluative techniques and concluded that the only approach which addresses itself to the critical issue of quality of care delivered is an assessment of clinical competence by an objective evaluation of a physician's performance in day-to-day practice utilizing peer review methods. We believe that this review should clearly display the physician's ability to obtain data in a logical and complete manner, and also to validate the clinical decisions derived from this data.

An awareness of the relatively primitive state of the art, and of several specific bottlenecks standing in the way of a viable competence evaluation caused the ASIM Board of Trustees to assign several committees specific tasks aimed at correction of these deficiencies:

We assigned high priority to the task of assuring adequate documentation of care in a manner which clearly displays the physician's ability to obtain data in a logical and complete manner and the validity of his clinical decisions. Believing that the problem-oriented approach to record-keeping fulfilled these qualifications, as well as offering other advantages over traditional methods, the Society has actively promoted the use of the Weed System among its membership.

We have developed a data collection system suitable for completion by a non-physician and also suitable for computerization. The system seeks to display in a "time-oriented" manner the initial data base, diagnostic, therapeutic services and continued care. This instrument has been subjected to test by members of the ASIM Board of Trustees and the Quality Evaluation Committee. Appropriate modifications have been made.

We have also developed critical process and outcome criteria for the majority of problems and diagnoses encountered by internists — these to be used for comparison with actual performance.

The ASIM Committee on Assessment by Performance (the Committee on Quality Evaluation name changed in May 1974) believes that progress has been made toward achievement of the above tasks and that the basic tools are now available for the analysis of the practice patterns of a limited population of internists. It is felt that refinement of these techniques and final determination of feasibility of approach awaits actual trial. For this reason, the Committee proposed, and the ASIM Board of Trustees approved earlier this year, that two demonstration projects be initiated by January 1, 1975, with the
following goal:

"Assessment of clinical competency by an objective evaluation of the physician's investigative, therapeutic and exploratory decisions based upon valid clinical data, and leading to an acceptable outcome. These decisions should be properly documented and arrived at efficiently."

The Model

The over-all conceptual model for the proposed assessment by performance process may be expressed as follows:

```
M.D. ACCEPTS ASSESS. BY PERF.  
SURVEY FOR PROFILE OF PRACTICE  
COMPARISON WITH CRITERIA AT DECISION POINTS
```

```
CONTINUING EDUCATION LOCAL PANEL  
AUDIT  
```

```
PASS  
```

```
FAIL  
```

```
RECERTIFY OR RELICENSE  
```

The model describes the general process of assessment by performance. This process begins with a physician agreeing to participate in the program and accepting the proposed methodology and results. The physician's practice is then surveyed to determine a profile of the practice, i.e., a frequency distribution of various problems and diagnoses that are treated by the physician. Certain patient charts are then abstracted for specific information. This information is compared with predetermined criteria that are related
to various stages in the clinical decision process. This comparison is concluded by a decision that the physician passes or fails the assessment process. Those who pass this stage of the assessment are automatically recertified, and those who fail are referred to continuing education programs, and are assigned another date for repetition of the screening/audit procedure.

This model, if successful, could provide a more meaningful alternative to examination as one approach to recertification or relicensure, whether voluntary or mandatory, and lend full credibility to our assurance to the public of high quality care. Also, if successful, this technique with its emphasis on logic branching evaluation and outcome analysis may open the door to significant practical advances in the art of quality assessment and consequently have considerable impact on PSRO and related activities.

Methodology

Two demonstration sites have been selected for this study. One will be centered at the New Haven, Connecticut area and the other in Columbus, Ohio. Daniel Hamaty, M.D., will coordinate the Connecticut demonstration and William Millhon, M.D., will coordinate the Ohio group. Each demonstration group will consist of 10 to 20 internists in private practice who are expected to volunteer their participation in the project.

The number of participants in each group was determined as follows. According to the AMA, there are approximately 30,000 internists in private practice who are either self-designated or board eligible or board certified in Internal Medicine. If all of these internists wish to participate in future assessment programs, it will be necessary to assess 3,000 internists every year if there is a 10 year interval between assessments. The assessment process is expected to require six months, thus two assessment periods can be designated in every year. The 3,000 internists to be assessed every year could be divided into two groups of 1,500 each. The assessment is expected to be conducted in a number of regions. Assuming approximately 200 regions (possibly related to the PSRO regions) the assessment will involve no more than 20 physicians per region per assessment period.

The Columbus, Ohio and the New Haven, Connecticut regions were selected because of the previous experience of Dr. Hamaty and because of strong leadership and commitment expressed by both the Connecticut and the Ohio groups.

Observation and Data Collection

We believe that the most economical method for data collection involves the utilization of a standardized form to be completed by the physician himself, his staff, or a trained health accountant. If either staff or a health accountant is used for the data collec-
tion, the physician must certify that the required information is available in the patient's chart. Each physician may train his own staff to collect the information or he may request the aid of a health accountant.

Health Accountants

The experience of Dr. Hamaty with the Connecticut Ambulatory Care Study of the Connecticut State Medical Society will be tapped for the purpose of training health accountants. Personnel of the Connecticut Ambulatory Care Project will train a team of health accountants for the Assessment by Performance Project in the Connecticut area. A medical record librarian or a registered nurse from Ohio will spend approximately two weeks of intensive training in Connecticut and will then be charged with the training of the Ohio team of health accountants. It is expected that three health accountants will be needed in each region.

The health accountants will have two major tasks: The first will be to review the participants' practices and determine a practice profile for each participant. The second task will be to review patient charts within each participant's practice and complete the testing instruments for designated patients in the most reliable and accurate way possible.

The Sample

In constructing the practice profiles for each participant, the health accountants will review the appointment books and select at random 20% of the appointments made on 35 different days, selected at random from a one year period. Patient charts will be retrieved for the selected cases to confirm the patients' diagnoses and complaints. The outcome of this phase will be a profile of each practice in the form of a distribution of the most common problems and diagnoses seen and treated by each participant.

Using these profiles as guides, the sample of patients' records to be abstracted in each practice will be selected at random in proportion to the respective profiles of each participant. No more than 100 patients will be reviewed per practice during a period of six months; however, the precise sample size needed will depend on the accuracy and validity of the testing instrument. At the present time, it is difficult to assess the error rate that can be expected to be generated by the use of the instrument itself. In addition to errors generated by the testing instrument, the classical error analysis must be considered.

Since we are dealing with a testing procedure, we must accept a certain error level. Two types of errors are usually recognized in this type of testing activity, the false negatives and the false positives. For statistical purposes, it is convenient to describe these two errors as follows: a) The Type I or Alpha error, which
specifies the fraction of good performers that may be rejected as poor ones; b) The Type II or Beta error, which specifies the fraction of cases of poor performance that may be accepted as good performance. These two types of errors are the main factors to be considered in the sampling of cases for review. In addition to these two parameters (to be designated by the Committee), it is also necessary to identify an acceptable level of poor performance in each practice. That is to say, we must accept the fact that as human beings we are all bound to commit some errors in judgment and even in the practice of the best clinician one may detect some cases of poor performance. This may be given as a percent of cases that one may identify as poor care within any practice without concluding that the clinician is a poor performer. In addition, it is necessary to identify that critical poor performance level.

'These parameters are not readily available and it is expected that the first few experimentations will yield some information that may assist in identifying these values.

The Testing Instrument

The objectives of the instrument are:

1. To aid in an objective evaluation of the quality, efficiency and outcome of the various steps in patient care.
2. To aid in assuring credibility for provider.
3. To provide uniformity in the methodology so that its credibility is maintained regardless of the practice setting in which it is used.

In developing this testing instrument, the following are some of the specifications which must be satisfied:

1. It should be problem oriented.
2. It must have a problem-diagnosis specificity. (i.e., it must relate to both problems and diagnoses and to the logical relationship between them)
3. It must be computer compatible for volume processing.
4. A representative sample of cases for each problem-diagnosis must be audited periodically.
5. The instrument must be adaptable to either retrospective or concurrent review.
6. It must reflect both inpatient and ambulatory care.
7. It must identify specific outcomes in the process of medical care.
8. It should be related to tracer problem-diagnosis sets, rather than attempting to deal with all problems and diagnoses that a physician may encounter.

The main premise of the testing instrument is that medical care should be a logical process consisting of several definable steps, each step ending in a predictable decision point. (see attachment #1). Each decision point can then be validated by an analysis of the data that has been accumulated prior to that point. Each decision point will then justify the process of medical care that follows. In the diagram, each decision point is a locus for the evaluation of the outcome of patient care given the information at that point. For example, at the end of history and physical, the evaluation may yield an identified problem that will lead the physician to order problem solving ancillary services. If, however, no problem has been identified, the physician will order problem seeking ancillary services. After the results of the ancillary information is available, a higher level of problem resolution or established diagnosis will result. This will lead the physician to decide whether to prescribe therapeutic management for the treatment of the specific diagnosis that has been established or to discharge the patient because he has been found to be in good health. Similarly, after the treatment has been instituted, an assessment of the outcome of care may lead the physician to decide to, (1) discharge the patient because the disease has been eliminated, (2) to change the treatment regime because of an unsatisfactory therapeutic outcome, (3) to progress to continuing care because the disease has been adequately controlled, (4) to re-evaluate the patient because of inadequate problem resolution.

The decision point criteria are being perfected by the ASIM Norms Committee. These criteria will be validated by specialists in each field. Two types of criteria can be identified:

1. **Critical criteria:** Those procedures or treatments that will be present with a 100% incidence except when specific exclusions are noted.

2. **Relevant criteria:** Those procedures and treatments that are indicated only when specific findings have been obtained during the investigative management, at which time relevant criteria become critical criteria.

Since physicians see problems and complaints which must be resolved into diagnoses, the Norms Committee devised an approach for identifying most of the possible diagnoses resulting from a certain complaint. The attached example [Attachment #2] is the possible resolution of the problem "chest pain."

The Norms Committee has identified a number of common problems and diseases in the practice of internal medicine for which they have developed quality care criteria. Attachment #3 is a sample of the form used by the Committee to develop a model treatment.
The ASIM Norms Committee, in conjunction with the Committee on Assessment by Performance, has developed a testing instrument that is based on the previously described philosophy and criteria. (Attachment #4)

The proposed pilot experiment will reveal the feasibility of using this testing instrument in clinical practice including the level of physician satisfaction.

Evaluation of the Assessment Data

The health accountants will be provided templates of critical criteria for specific problems and diagnoses. These templates will be in the form of an overlay to fit over the data collection form. They will be trained to retrieve from the patient chart the specific items required by each template. First evaluation will involve a simple count of all critical criteria to assure 100% compliance. Further analysis must evaluate the appropriateness of decisions as related to previously collected information. Final analysis will be made of outcomes predicted and achieved.

Data Processing Requirements

It is planned to do most of the processing of the data from the demonstration project manually, not involving a computer. This is necessary to establish all the evaluation routines and identify all possible pitfalls. It is, however, our plan to engage a computer processing specialist to advise on hardware and software requirements for large processing of this information. Assuming that the demonstration will result in a feasible approach to assessment by performance, it may be necessary to use a computer for the screening of a large volume of cases that must be assessed during every testing period. It is expected that in full operation the system may be required to review approximately 30,000 patient records per year. The design of the testing instrument and the evaluation routines is compatible with computer processing and it is therefore expected that this task may be accomplished efficiently by a computer.

Proposed Time-table

The pilot experiment for the testing instruments began as scheduled on June 15, 1974. Evaluation and modifications of this testing instrument, along with completion of the critical criteria lists have been completed. Training of health accountants could be accomplished within six weeks. We are hoping to initiate the two demonstration programs by January 1975, proceeding with collection of data and evaluation until June 1975. Evaluation of the total program and preparation of the final reports should be completed by December 31, 1975.
Cost Control

The project will closely monitor the costs of the assessment process. Research costs will be separated from the costs of assessment efforts. It is expected that as a result of this careful cost analysis it may be possible to determine the cost per assessed physician. This information is needed for the establishment of a future charge per assessment. This may be the best indication for the feasibility and practicality of the proposed assessment mechanism because it will show whether or not such an assessment program can become self-sufficient.

If this proposed research demonstrates that the assessment procedure is feasible, applicants may be charged a fee for recertification. It is expected that these fees will cover the costs of the program, making it independent and self-sufficient.

Consultation

Consultation from recognized authorities in the field of medical care quality evaluation will be in the final developmental phase of the project and later in confirmation of the validity of results. The academic expertise of such possible consultants as Drs. John Williamson, Beverly Paine or Paul Sanazaro will be particularly valuable in complementing the clinical experience of the Committee.
ISCHEMIC HEART DISEASE, ANGINA PECTORIS

Tests & Procedures:

ECG
Post prandial blood sugar
or glucose tolerance test
Lipid profile

Stress test

Holter monitor

X-rays:

Chest x-ray
Coronary angiography

ECG
Critical, no exceptions
Post prandial blood sugar
Critical, no exceptions
or glucose tolerance test
Lipid profile

Stress test

Holter monitor

1) suspicion of dysrhythmia

X-rays:

Chest x-ray
Critical, no exceptions
Coronary angiography

1) for consideration of surgery
2) suspicion of ischemic heart disease

Management:

Nitroglycerine or analogue
Critical, no exceptions
Weight control
Relevant, if obesity is present
Beta adrenergic blocking agents
Relevant:
1) uncontrolled angina pectoris in the
    absence of congestive heart failure or asthma
Exercise program
Relevant, after multistage exercise testing has established tolerance
<table>
<thead>
<tr>
<th>ATTACHMENT NO.</th>
<th>ATTACH. DESCR.</th>
<th>NAME</th>
<th>TITLE</th>
<th>DATE</th>
<th>PLACE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>